AWS Data Pipeline

API Reference

API Version 2012-10-29

AWS Data Pipeline: API Reference

Table of Contents

Welcome

AWS Data Pipeline is a web service that configures and manages a data-driven workflow called a pipeline. AWS Data Pipeline handles the details of scheduling and ensuring that data dependencies are met so your application can focus on processing the data.

The AWS Data Pipeline SDKs and CLI implements two main sets of functionality. The first set of actions configure the pipeline in the web service. You perform these actions to create a pipeline and define data sources, schedules, dependencies, and the transforms to be performed on the data.

The second set of actions are used by a task runner application that calls the AWS Data Pipeline service to receive the next task ready for processing. The logic for performing the task, such as querying the data, running data analysis, or converting the data from one format to another, is contained within the task runner. The task runner performs the task assigned to it by the web service, reporting progress to the web service as it does so. When the task is done, the task runner reports the final success or failure of the task to the web service.

AWS Data Pipeline provides a JAR implementation of a task runner called AWS Data Pipeline Task Runner. AWS Data Pipeline Task Runner provides logic for common data management scenarios, such as performing database queries and running data analysis using Amazon Elastic MapReduce (Amazon EMR). You can use AWS Data Pipeline Task Runner as your task runner, or you can write your own task runner to provide custom data management.

This document was last updated on November 19, 2014.

Actions

The following actions are supported:

ActivatePipeline

Validates a pipeline and initiates processing. If the pipeline does not pass validation, activation fails. You cannot perform this operation on FINISHED pipelines and attempting to do so will return an InvalidRequestException.

Call this action to start processing pipeline tasks of a pipeline you've created using the CreatePipeline (p. 5) and PutPipelineDefinition (p. 31) actions. A pipeline cannot be modified after it has been successfully activated.

Request Syntax

```
{
    "ParameterValues": [
        {
            "Id": "string",
            "StringValue": "string"
        }
    ],
    "PipelineId": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

ParameterValues
    Returns a list of parameter values to pass to the pipeline at activation.

    Type: array of ParameterValue (p. 66) objects

    Required: No
PipelineId
    The identifier of the pipeline to activate.

    Type: String

    Length constraints: Minimum length of 1. Maximum length of 1024.

    Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError

An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException

The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

PipelineDeletedException

The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException

The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.ActivatePipeline
Content-Length: 39
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-06372391ZG65EXAMPLE"}
```

Sample Response

```
HTTP/1.1 200
x-amzn-RequestId: ee19d5bf-074e-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 2
Date: Mon, 12 Nov 2012 17:50:53 GMT

{}
```

CreatePipeline

Creates a new empty pipeline. When this action succeeds, you can then use the PutPipelineDefinition (p. 31) action to populate the pipeline.

Request Syntax

```
{
    "Description": "string",
    "Name": "string",
    "Tags": [
        {
            "Key": "string",
            "Value": "string"
        }
    ],
    "UniqueId": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

Description
The description of the new pipeline.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

Name
The name of the new pipeline. You can use the same name for multiple pipelines associated with your AWS account, because AWS Data Pipeline assigns each new pipeline a unique pipeline identifier.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

Tags
Type: array of Tag (p. 71) objects

Length constraints: Minimum of 0 item(s) in the list. Maximum of 10 item(s) in the list.

Required: No

UniqueId
A unique identifier that you specify. This identifier is not the same as the pipeline identifier assigned by AWS Data Pipeline. You are responsible for defining the format and ensuring the uniqueness of this identifier. You use this parameter to ensure idempotency during repeated calls to CreateP-

ipeline (p. 5). For example, if the first call to CreatePipeline (p. 5) does not return a clear success, you can pass in the same unique identifier and pipeline name combination on a subsequent call to CreatePipeline (p. 5). CreatePipeline (p. 5) ensures that if a pipeline already exists with the same name and unique identifier, a new pipeline will not be created. Instead, you'll receive the pipeline identifier from the previous attempt. The uniqueness of the name and unique identifier combination is scoped to the AWS account or IAM user credentials.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

Response Syntax

```
{
    "PipelineId": "string"
}
```

Response Elements

The following data is returned in JSON format by the service.

PipelineId
The ID that AWS Data Pipeline assigns the newly created pipeline. The ID is a string of the form: df-06372391ZG65EXAMPLE.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException
The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.CreatePipeline
Content-Length: 91
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"name": "myPipeline",
 "uniqueId": "123456789",
 "description": "This is my first pipeline"}
```

Sample Response

```
HTTP/1.1 200
x-amzn-RequestId: b16911ce-0774-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 40
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"pipelineId": "df-06372391ZG65EXAMPLE"}
```

DeletePipeline

Permanently deletes a pipeline, its pipeline definition and its run history. You cannot query or restore a deleted pipeline. AWS Data Pipeline will attempt to cancel instances associated with the pipeline that are currently being processed by task runners. Deleting a pipeline cannot be undone.

To temporarily pause a pipeline instead of deleting it, call SetStatus (p. 47) with the status set to Pause on individual components. Components that are paused by SetStatus (p. 47) can be resumed.

Request Syntax

```
{
    "PipelineId": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

PipelineId
    The identifier of the pipeline to be deleted.

    Type: String

    Length constraints: Minimum length of 1. Maximum length of 1024.

    Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
    An internal service error occurred.

    HTTP Status Code: 500

InvalidRequestException
    The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

    HTTP Status Code: 400

PipelineNotFoundException
    The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.DeletePipeline
Content-Length: 50
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-06372391ZG65EXAMPLE"}
```

Sample Response

```
x-amzn-RequestId: b7a88c81-0754-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 0
Date: Mon, 12 Nov 2012 17:50:53 GMT

Unexpected response: 200, OK, undefined
```

DescribeObjects

Returns the object definitions for a set of objects associated with the pipeline. Object definitions are composed of a set of fields that define the properties of the object.

Request Syntax

```
{
    "EvaluateExpressions": "boolean",
    "Marker": "string",
    "ObjectIds": [
        "string"
    ],
    "PipelineId": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

EvaluateExpressions
> Indicates whether any expressions in the object should be evaluated when the object descriptions are returned.
>
> Type: Boolean
>
> Required: No

Marker
> The starting point for the results to be returned. The first time you call DescribeObjects (p. 10), this value should be empty. As long as the action returns `HasMoreResults` as `True`, you can call DescribeObjects (p. 10) again and pass the marker value from the response to retrieve the next set of results.
>
> Type: String
>
> Length constraints: Minimum length of 0. Maximum length of 1024.
>
> Required: No

ObjectIds
> Identifiers of the pipeline objects that contain the definitions to be described. You can pass as many as 25 identifiers in a single call to DescribeObjects.
>
> Type: array of Strings
>
> Required: Yes

PipelineId
> Identifier of the pipeline that contains the object definitions.
>
> Type: String
>
> Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

Response Syntax

```
{
    "HasMoreResults": "boolean",
    "Marker": "string",
    "PipelineObjects": [
        {
            "Fields": [
                {
                    "Key": "string",
                    "RefValue": "string",
                    "StringValue": "string"
                }
            ],
            "Id": "string",
            "Name": "string"
        }
    ]
}
```

Response Elements

The following data is returned in JSON format by the service.

HasMoreResults
If `True`, there are more pages of results to return.

Type: Boolean

Marker
The starting point for the next page of results. To view the next page of results, call DescribeObjects (p. 10) again with this marker value.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

PipelineObjects
An array of object definitions that are returned by the call to DescribeObjects (p. 10).

Type: array of PipelineObject (p. 68) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException

The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

PipelineDeletedException

The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException

The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.DescribeObjects
Content-Length: 98
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-06372391ZG65EXAMPLE",
 "objectIds":
  ["Schedule"],
 "evaluateExpressions": true}
```

Sample Response

```
x-amzn-RequestId: 4c18ea5d-0777-11e2-8a14-21bb8a1f50ef
Content-Type: application/x-amz-json-1.1
Content-Length: 1488
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"hasMoreResults": false,
 "pipelineObjects":
  [
    {"fields":
      [
        {"key": "startDateTime",
         "stringValue": "2012-12-12T00:00:00"},
        {"key": "parent",
         "refValue": "Default"},
```

```
        {"key": "@sphere",
         "stringValue": "COMPONENT"},
        {"key": "type",
         "stringValue": "Schedule"},
        {"key": "period",
         "stringValue": "1 hour"},
        {"key": "endDateTime",
         "stringValue": "2012-12-21T18:00:00"},
        {"key": "@version",
         "stringValue": "1"},
        {"key": "@status",
         "stringValue": "PENDING"},
        {"key": "@pipelineId",
         "stringValue": "df-06372391ZG65EXAMPLE"}
      ],
     "id": "Schedule",
     "name": "Schedule"}
   ]
}
```

DescribePipelines

Retrieve metadata about one or more pipelines. The information retrieved includes the name of the pipeline, the pipeline identifier, its current state, and the user account that owns the pipeline. Using account credentials, you can retrieve metadata about pipelines that you or your IAM users have created. If you are using an IAM user account, you can retrieve metadata about only those pipelines you have read permission for.

To retrieve the full pipeline definition instead of metadata about the pipeline, call the GetPipelineDefinition (p. 20) action.

Request Syntax

```
{
    "PipelineIds": [
        "string"
    ]
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

PipelineIds

Identifiers of the pipelines to describe. You can pass as many as 25 identifiers in a single call to DescribePipelines (p. 14). You can obtain pipeline identifiers by calling ListPipelines (p. 24).

Type: array of Strings

Required: Yes

Response Syntax

```
{
    "PipelineDescriptionList": [
        {
            "Description": "string",
            "Fields": [
                {
                    "Key": "string",
                    "RefValue": "string",
                    "StringValue": "string"
                }
            ],
            "Name": "string",
            "PipelineId": "string",
            "Tags": [
                {
```

```
                    "Key": "string",
                    "Value": "string"
                }
            ]
        }
    ]
}
```

Response Elements

The following data is returned in JSON format by the service.

PipelineDescriptionList
An array of descriptions returned for the specified pipelines.

Type: array of PipelineDescription (p. 66) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException
The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

PipelineDeletedException
The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException
The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.DescribePipelines
Content-Length: 70
Host: datapipeline.us-east-1.amazonaws.com
```

```
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineIds":
  ["df-08785951KAKJEXAMPLE"]
}
```

Sample Response

```
x-amzn-RequestId: 02870eb7-0736-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 767
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"pipelineDescriptionList":
  [
    {"description": "This is my first pipeline",
     "fields":
      [
        {"key": "@pipelineState",
         "stringValue": "SCHEDULED"},
        {"key": "description",
         "stringValue": "This is my first pipeline"},
        {"key": "name",
         "stringValue": "myPipeline"},
        {"key": "@creationTime",
         "stringValue": "2012-12-13T01:24:06"},
        {"key": "@id",
         "stringValue": "df-0937003356ZJEXAMPLE"},
        {"key": "@sphere",
         "stringValue": "PIPELINE"},
        {"key": "@version",
         "stringValue": "1"},
        {"key": "@userId",
         "stringValue": "924374875933"},
        {"key": "@accountId",
         "stringValue": "924374875933"},
        {"key": "uniqueId",
         "stringValue": "1234567890"}
      ],
     "name": "myPipeline",
     "pipelineId": "df-0937003356ZJEXAMPLE"}
  ]
}
```

EvaluateExpression

Evaluates a string in the context of a specified object. A task runner can use this action to evaluate SQL queries stored in Amazon S3.

Request Syntax

```
{
    "Expression": "string",
    "ObjectId": "string",
    "PipelineId": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

Expression
The expression to evaluate.

Type: String

Length constraints: Minimum length of 0. Maximum length of 20971520.

Required: Yes

ObjectId
The identifier of the object.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

PipelineId
The identifier of the pipeline.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

Response Syntax

```
{
    "EvaluatedExpression": "string"
}
```

Response Elements

The following data is returned in JSON format by the service.

EvaluatedExpression
The evaluated expression.

Type: String

Length constraints: Minimum length of 0. Maximum length of 20971520.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException
The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

PipelineDeletedException
The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException
The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

TaskNotFoundException
The specified task was not found.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.DescribePipelines
Content-Length: 164
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-08785951KAKJEXAMPLE",
```

```
        "objectId": "Schedule",
        "expression": "Transform started at #{startDateTime} and finished at
#{endDateTime}"}
```

Sample Response

```
x-amzn-RequestId: 02870eb7-0736-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 103
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"evaluatedExpression": "Transform started at 2012-12-12T00:00:00 and finished
 at 2012-12-21T18:00:00"}
```

GetPipelineDefinition

Returns the definition of the specified pipeline. You can call GetPipelineDefinition (p. 20) to retrieve the pipeline definition you provided using PutPipelineDefinition (p. 31).

Request Syntax

```
{
    "PipelineId": "string",
    "Version": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

PipelineId
The identifier of the pipeline.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

Version
The version of the pipeline definition to retrieve. This parameter accepts the values `latest` (default) and `active`. Where `latest` indicates the last definition saved to the pipeline and `active` indicates the last definition of the pipeline that was activated.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

Response Syntax

```
{
    "ParameterObjects": [
        {
            "Attributes": [
                {
                    "Key": "string",
                    "StringValue": "string"
                }
            ],
            "Id": "string"
        }
```

```
        ],
    "ParameterValues": [
        {
            "Id": "string",
            "StringValue": "string"
        }
    ],
    "PipelineObjects": [
        {
            "Fields": [
                {
                    "Key": "string",
                    "RefValue": "string",
                    "StringValue": "string"
                }
            ],
            "Id": "string",
            "Name": "string"
        }
    ]
}
```

Response Elements

The following data is returned in JSON format by the service.

ParameterObjects
>   Returns a list of parameter objects used in the pipeline definition.
>
>   Type: array of ParameterObject (p. 65) objects

ParameterValues
>   Returns a list of parameter values used in the pipeline definition.
>
>   Type: array of ParameterValue (p. 66) objects

PipelineObjects
>   An array of objects defined in the pipeline.
>
>   Type: array of PipelineObject (p. 68) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
>   An internal service error occurred.
>
>   HTTP Status Code: 500

InvalidRequestException
>   The request was not valid. Verify that your request was properly formatted, that the signature was
>   generated with the correct credentials, and that you haven't exceeded any of the service limits for
>   your account.
>
>   HTTP Status Code: 400

PipelineDeletedException

> The specified pipeline has been deleted.

> HTTP Status Code: 400

PipelineNotFoundException

> The specified pipeline was not found. Verify that you used the correct user and account identifiers.

> HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.GetPipelineDefinition
Content-Length: 40
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams


{"pipelineId": "df-06372391ZG65EXAMPLE"}
```

Sample Response

```
x-amzn-RequestId: e28309e5-0776-11e2-8a14-21bb8a1f50ef
Content-Type: application/x-amz-json-1.1
Content-Length: 890
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"pipelineObjects":
  [
    {"fields":
      [
        {"key": "workerGroup",
         "stringValue": "workerGroup"}
      ],
     "id": "Default",
     "name": "Default"},
    {"fields":
      [
        {"key": "startDateTime",
         "stringValue": "2012-09-25T17:00:00"},
        {"key": "type",
         "stringValue": "Schedule"},
        {"key": "period",
         "stringValue": "1 hour"},
```

```
        {"key": "endDateTime",
         "stringValue": "2012-09-25T18:00:00"}
      ],
     "id": "Schedule",
     "name": "Schedule"},
    {"fields":
      [
        {"key": "schedule",
         "refValue": "Schedule"},
        {"key": "command",
         "stringValue": "echo hello"},
        {"key": "parent",
         "refValue": "Default"},
        {"key": "type",
         "stringValue": "ShellCommandActivity"}
      ],
     "id": "SayHello",
     "name": "SayHello"}
  ]
}
```

ListPipelines

Returns a list of pipeline identifiers for all active pipelines. Identifiers are returned only for pipelines you have permission to access.

Request Syntax

```
{
    "Marker": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

Marker
The starting point for the results to be returned. The first time you call ListPipelines (p. 24), this value should be empty. As long as the action returns HasMoreResults as True, you can call List-Pipelines (p. 24) again and pass the marker value from the response to retrieve the next set of results.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

Response Syntax

```
{
    "HasMoreResults": "boolean",
    "Marker": "string",
    "PipelineIdList": [
        {
            "Id": "string",
            "Name": "string"
        }
    ]
}
```

Response Elements

The following data is returned in JSON format by the service.

HasMoreResults
If True, there are more results that can be obtained by a subsequent call to ListPipelines (p. 24).

Type: Boolean

Marker

If not null, indicates the starting point for the set of pipeline identifiers that the next call to List-Pipelines (p. 24) will retrieve. If null, there are no more pipeline identifiers.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

PipelineIdList

A list of all the pipeline identifiers that your account has permission to access. If you require additional information about the pipelines, you can use these identifiers to call DescribePipelines (p. 14) and GetPipelineDefinition (p. 20).

Type: array of PipelineIdName (p. 67) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError

An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException

The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.ListPipelines
Content-Length: 14
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{}
```

Sample Response

```
Status:
x-amzn-RequestId: b3104dc5-0734-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 39
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"PipelineIdList":
  [
    {"id": "df-08785951KAKJEXAMPLE",
     "name": "MyPipeline"},
    {"id": "df-08662578ISYEXAMPLE",
     "name": "MySecondPipeline"}
  ]
}
```

PollForTask

Task runners call this action to receive a task to perform from AWS Data Pipeline. The task runner specifies which tasks it can perform by setting a value for the workerGroup parameter of the PollForTask (p. 27) call. The task returned by PollForTask (p. 27) may come from any of the pipelines that match the worker-Group value passed in by the task runner and that was launched using the IAM user credentials specified by the task runner.

If tasks are ready in the work queue, PollForTask (p. 27) returns a response immediately. If no tasks are available in the queue, PollForTask (p. 27) uses long-polling and holds on to a poll connection for up to a 90 seconds during which time the first newly scheduled task is handed to the task runner. To accomodate this, set the socket timeout in your task runner to 90 seconds. The task runner should not call Poll-ForTask (p. 27) again on the same `workerGroup` until it receives a response, and this may take up to 90 seconds.

Request Syntax

```
{
    "Hostname": "string",
    "InstanceIdentity": {
        "Document": "string",
        "Signature": "string"
    },
    "WorkerGroup": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

Hostname
> The public DNS name of the calling task runner.
>
> Type: String
>
> Length constraints: Minimum length of 1. Maximum length of 1024.
>
> Required: No

InstanceIdentity
> Identity information for the Amazon EC2 instance that is hosting the task runner. You can get this value by calling the URI, `http://169.254.169.254/latest/meta-data/instance-id`, from the EC2 instance. For more information, go to Instance Metadata in the *Amazon Elastic Compute Cloud User Guide.* Passing in this value proves that your task runner is running on an EC2 instance, and ensures the proper AWS Data Pipeline service charges are applied to your pipeline.
>
> Type: InstanceIdentity (p. 63) object
>
> Required: No

WorkerGroup
> Indicates the type of task the task runner is configured to accept and process. The worker group is set as a field on objects in the pipeline when they are created. You can only specify a single value

for `workerGroup` in the call to PollForTask (p. 27). There are no wildcard values permitted in `workerGroup`, the string must be an exact, case-sensitive, match.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: Yes

Response Syntax

```
{
    "TaskObject": {
        "AttemptId": "string",
        "Objects":
            {
                "string" :
                    {
                        "Fields": [
                            {
                                "Key": "string",
                                "RefValue": "string",
                                "StringValue": "string"
                            }
                        ],
                        "Id": "string",
                        "Name": "string"
                    }
            },
        "PipelineId": "string",
        "TaskId": "string"
    }
}
```

Response Elements

The following data is returned in JSON format by the service.

TaskObject
> An instance of PollForTaskResult (p. 68), which contains an instance of TaskObject (p. 72). The returned object contains all the information needed to complete the task that is being assigned to the task runner. One of the fields returned in this object is taskId, which contains an identifier for the task being assigned. The calling task runner uses taskId in subsequent calls to ReportTaskProgress (p. 41) and SetTaskStatus (p. 50).

> Type: TaskObject (p. 72) object

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError

An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException

The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

TaskNotFoundException

The specified task was not found.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.PollForTask
Content-Length: 59
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"workerGroup": "MyworkerGroup",
 "hostname": "example.com"}
```

Sample Response

```
x-amzn-RequestId: 41c713d2-0775-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 39
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"taskObject":
  {"attemptId": "@SayHello_2012-12-12T00:00:00_Attempt=1",
   "objects":
    {"@SayHello_2012-12-12T00:00:00_Attempt=1":
      {"fields":
        [
          {"key": "@componentParent",
           "refValue": "SayHello"},
          {"key": "@scheduledStartTime",
           "stringValue": "2012-12-12T00:00:00"},
          {"key": "parent",
```

```
                "refValue": "SayHello"},
            {"key": "@sphere",
             "stringValue": "ATTEMPT"},
            {"key": "workerGroup",
             "stringValue": "workerGroup"},
            {"key": "@instanceParent",
             "refValue": "@SayHello_2012-12-12T00:00:00"},
            {"key": "type",
             "stringValue": "ShellCommandActivity"},
            {"key": "@status",
             "stringValue": "WAITING_FOR_RUNNER"},
            {"key": "@version",
             "stringValue": "1"},
            {"key": "schedule",
             "refValue": "Schedule"},
            {"key": "@actualStartTime",
             "stringValue": "2012-12-13T01:40:50"},
            {"key": "command",
             "stringValue": "echo hello"},
            {"key": "@scheduledEndTime",
             "stringValue": "2012-12-12T01:00:00"},
            {"key": "@activeInstances",
             "refValue": "@SayHello_2012-12-12T00:00:00"},
            {"key": "@pipelineId",
             "stringValue": "df-0937003356ZJEXAMPLE"}
        ],
        "id": "@SayHello_2012-12-12T00:00:00_Attempt=1",
        "name": "@SayHello_2012-12-12T00:00:00_Attempt=1"}
    },
    "pipelineId": "df-0937003356ZJEXAMPLE",
    "taskId": "2xaM4wRs5zOsIH+g9U3oVHfAgAlbSqU6XduncB0HhZ3xMnmvfePZPn4dIbYXHy
WyRK+cU15MqDHwdrvftx/4wv+sNS4w34vJfv7QA9aOoOazW28l1GYSb2ZRR0N0paiQp+d1MhSKo10hOT
WOsVK5S5Lnx9Qm6omFgXHyIvZRIvTlrQMpr1xuUrflyGOfbFOGpOLpvPE172MYdqpZKnbSS4TcuqgQK
SWV2833fEubI57DPOP7ghWa2TcYeSIv4pdLYG53fTuwfbnbdc98g2LNUQzSVhSnt7BoqyN
wht2aQ6b/UHg9A80+KVpuXuqmz3m1MXwHFgxjdmuesXNOrrlGpeL
CcRWD+aGo0RN1NqhQRzNAig8V4GlaPTQzMsRCljKqvrIyAoP3Tt2XEGsHkkQo12rEX8Z90957XX2qKR
whruwYzqGkSLWjINoLdAxUJdpRXRc5DJTrBd3D5mdzn7kY1l7NEh4kFHJDt3Cx4Z3Mk8MYCA
CyCk/CEyy9DwuPi66cLz0NBcgbCM5LKjTBOwo1m+am+pvM1kSposE9FPP1+RFGb8k6jQB
TJx3TRz1yKilnGXQTZ5xvdOFpJrklIT0OXP1MG3+auM9FlJA+1dX90QoNJE5z7axmK//MO
GXUdkqFe2kiDkorqjxwDvc0Js9pVKfKvAmW8YqUbmI9l0ERpWCXXnLVHNmPWz3jaPY+OBAmuJWD
mxB/Z8p94aEDg4BVXQ7LvsKQ3DLYhaB7yJ390CJT+i0mm+EBqY60V6YikPSWD
FrYQ/NPi2b1DgE19mX8zHqw8qprIl4yh1Ckx2Iige4En/N5ktOoIxnASxAw/TzcE2skxdw5Kl
HDF+UTj71m16CR/dIaKlXijlfNlNzUBo/bNSadCQn3G5NoO501wPKI:XO50TgDNyo8EXAMPLE/g==:1"}
}
```

PutPipelineDefinition

Adds tasks, schedules, and preconditions that control the behavior of the pipeline. You can use Put-PipelineDefinition to populate a new pipeline.

PutPipelineDefinition (p. 31) also validates the configuration as it adds it to the pipeline. Changes to the pipeline are saved unless one of the following three validation errors exists in the pipeline.

1. An object is missing a name or identifier field.
2. A string or reference field is empty.
3. The number of objects in the pipeline exceeds the maximum allowed objects.
4. The pipeline is in a FINISHED state.

Pipeline object definitions are passed to the PutPipelineDefinition (p. 31) action and returned by the GetPipelineDefinition (p. 20) action.

Request Syntax

```
{
    "ParameterObjects": [
        {
            "Attributes": [
                {
                    "Key": "string",
                    "StringValue": "string"
                }
            ],
            "Id": "string"
        }
    ],
    "ParameterValues": [
        {
            "Id": "string",
            "StringValue": "string"
        }
    ],
    "PipelineId": "string",
    "PipelineObjects": [
        {
            "Fields": [
                {
                    "Key": "string",
                    "RefValue": "string",
                    "StringValue": "string"
                }
            ],
            "Id": "string",
            "Name": "string"
        }
    ]
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

ParameterObjects
  A list of parameter objects used with the pipeline.

  Type: array of ParameterObject (p. 65) objects

  Required: No
ParameterValues
  A list of parameter values used with the pipeline.

  Type: array of ParameterValue (p. 66) objects

  Required: No
PipelineId
  The identifier of the pipeline to be configured.

  Type: String

  Length constraints: Minimum length of 1. Maximum length of 1024.

  Required: Yes
PipelineObjects
  The objects that define the pipeline. These will overwrite the existing pipeline definition.

  Type: array of PipelineObject (p. 68) objects

  Required: Yes

Response Syntax

```
{
    "Errored": "boolean",
    "ValidationErrors": [
        {
            "Errors": [
                "string"
            ],
            "Id": "string"
        }
    ],
    "ValidationWarnings": [
        {
            "Id": "string",
            "Warnings": [
                "string"
            ]
        }
    ]
}
```

Response Elements

The following data is returned in JSON format by the service.

Errored

If `True`, there were validation errors. If errored is `True`, the pipeline definition is stored but cannot be activated until you correct the pipeline and call PutPipelineDefinition (p. 31) to commit the corrected pipeline.

Type: Boolean

ValidationErrors

A list of the validation errors that are associated with the objects defined in `pipelineObjects`.

Type: array of ValidationError (p. 73) objects

ValidationWarnings

A list of the validation warnings that are associated with the objects defined in `pipelineObjects`.

Type: array of ValidationWarning (p. 74) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError

An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException

The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

PipelineDeletedException

The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException

The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

Examples

Example 1

This example sets an valid pipeline configuration and returns success.

Sample Request

```
POST / HTTP/1.1
```

```
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.PutPipelineDefinition
Content-Length: 914
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-0937003356ZJEXAMPLE",
 "pipelineObjects":
  [
    {"id": "Default",
      "name": "Default",
      "fields":
       [
          {"key": "workerGroup",
           "stringValue": "workerGroup"}
       ]
    },
    {"id": "Schedule",
     "name": "Schedule",
      "fields":
       [
        {"key": "startDateTime",
          "stringValue": "2012-12-12T00:00:00"},
         {"key": "type",
          "stringValue": "Schedule"},
         {"key": "period",
          "stringValue": "1 hour"},
         {"key": "endDateTime",
          "stringValue": "2012-12-21T18:00:00"}
       ]
    },
    {"id": "SayHello",
     "name": "SayHello",
      "fields":
       [
          {"key": "type",
           "stringValue": "ShellCommandActivity"},
          {"key": "command",
           "stringValue": "echo hello"},
          {"key": "parent",
           "refValue": "Default"},
          {"key": "schedule",
           "refValue": "Schedule"}
       ]
    }
  ]
}
```

Sample Response

```
HTTP/1.1 200
```

```
x-amzn-RequestId: f74afc14-0754-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 18
Date: Mon, 12 Nov 2012 17:50:53 GMT


{"errored": false}
```

Example 2

This example sets an invalid pipeline configuration (the value for workerGroup is an empty string) and returns an error message.

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.PutPipelineDefinition
Content-Length: 903
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-06372391ZG65EXAMPLE",
 "pipelineObjects":
  [
    {"id": "Default",
     "name": "Default",
     "fields":
      [
        {"key": "workerGroup",
         "stringValue": ""}
      ]
    },
    {"id": "Schedule",
     "name": "Schedule",
     "fields":
      [
        {"key": "startDateTime",
         "stringValue": "2012-09-25T17:00:00"},
        {"key": "type",
         "stringValue": "Schedule"},
        {"key": "period",
         "stringValue": "1 hour"},
        {"key": "endDateTime",
         "stringValue": "2012-09-25T18:00:00"}
      ]
    },
    {"id": "SayHello",
     "name": "SayHello",
     "fields":
      [
        {"key": "type",
```

```
        "stringValue": "ShellCommandActivity"},
      {"key": "command",
       "stringValue": "echo hello"},
      {"key": "parent",
       "refValue": "Default"},
      {"key": "schedule",
       "refValue": "Schedule"}

    ]
  }
 ]
}
```

Sample Response

```
HTTP/1.1 200
x-amzn-RequestId: f74afc14-0754-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 18
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"__type": "com.amazon.setl.webservice#InvalidRequestException",
 "message": "Pipeline definition has errors: Could not save the pipeline
definition due to FATAL errors: [com.amazon.setl.webservice.ValidationErr
or@108d7ea9] Please call Validate to validate your pipeline"}
```

QueryObjects

Queries a pipeline for the names of objects that match a specified set of conditions.

The objects returned by QueryObjects (p. 37) are paginated and then filtered by the value you set for query. This means the action may return an empty result set with a value set for marker. If `HasMoreResults` is set to `True`, you should continue to call QueryObjects (p. 37), passing in the returned value for marker, until `HasMoreResults` returns `False`.

Request Syntax

```
{
    "Limit": "number",
    "Marker": "string",
    "PipelineId": "string",
    "Query": {
        "Selectors": [
            {
                "FieldName": "string",
                "Operator": {
                    "Type": "string",
                    "Values": [
                        "string"
                    ]
                }
            }
        ]
    },
    "Sphere": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

Limit

Specifies the maximum number of object names that QueryObjects (p. 37) will return in a single call. The default value is 100.

Type: Number

Required: No

Marker

The starting point for the results to be returned. The first time you call QueryObjects (p. 37), this value should be empty. As long as the action returns `HasMoreResults` as `True`, you can call QueryObjects (p. 37) again and pass the marker value from the response to retrieve the next set of results.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

PipelineId

Identifier of the pipeline to be queried for object names.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

Query

Query that defines the objects to be returned. The Query (p. 69) object can contain a maximum of ten selectors. The conditions in the query are limited to top-level String fields in the object. These filters can be applied to components, instances, and attempts.

Type: Query (p. 69) object

Required: No

Sphere

Specifies whether the query applies to components or instances. Allowable values: COMPONENT, INSTANCE, ATTEMPT.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: Yes

Response Syntax

```
{
    "HasMoreResults": "boolean",
    "Ids": [
        "string"
    ],
    "Marker": "string"
}
```

Response Elements

The following data is returned in JSON format by the service.

HasMoreResults

If True, there are more results that can be obtained by a subsequent call to QueryObjects (p. 37).

Type: Boolean

Ids

A list of identifiers that match the query selectors.

Type: array of Strings

Marker

The starting point for the results to be returned. As long as the action returns HasMoreResults as True, you can call QueryObjects (p. 37) again and pass the marker value from the response to retrieve the next set of results.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException
The request was not valid. Verify that your request was properly formatted, that the signature was
generated with the correct credentials, and that you haven't exceeded any of the service limits for
your account.

HTTP Status Code: 400

PipelineDeletedException
The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException
The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.QueryObjects
Content-Length: 123
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-06372391ZG65EXAMPLE",
 "query":
  {"selectors":
    [
    ]
  },
 "sphere": "INSTANCE",
 "marker": "",
 "limit": 10}
```

Sample Response

```
x-amzn-RequestId: 14d704c1-0775-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 72
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"hasMoreResults": false,
 "ids":
   ["@SayHello_1_2012-09-25T17:00:00"]
}
```

ReportTaskProgress

Updates the AWS Data Pipeline service on the progress of the calling task runner. When the task runner is assigned a task, it should call ReportTaskProgress to acknowledge that it has the task within 2 minutes. If the web service does not recieve this acknowledgement within the 2 minute window, it will assign the task in a subsequent PollForTask (p. 27) call. After this initial acknowledgement, the task runner only needs to report progress every 15 minutes to maintain its ownership of the task. You can change this reporting time from 15 minutes by specifying a `reportProgressTimeout` field in your pipeline. If a task runner does not report its status after 5 minutes, AWS Data Pipeline will assume that the task runner is unable to process the task and will reassign the task in a subsequent response to PollForTask (p. 27). task runners should call ReportTaskProgress (p. 41) every 60 seconds.

Request Syntax

```
{
    "Fields": [
        {
            "Key": "string",
            "RefValue": "string",
            "StringValue": "string"
        }
    ],
    "TaskId": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

Fields
Key-value pairs that define the properties of the ReportTaskProgressInput object.

Type: array of Field (p. 62) objects

Required: No

TaskId
Identifier of the task assigned to the task runner. This value is provided in the TaskObject (p. 72) that the service returns with the response for the PollForTask (p. 27) action.

Type: String

Length constraints: Minimum length of 1. Maximum length of 2048.

Required: Yes

Response Syntax

```
{
```

```
      "Canceled": "boolean"
}
```

Response Elements

The following data is returned in JSON format by the service.

Canceled
If `True`, the calling task runner should cancel processing of the task. The task runner does not need to call SetTaskStatus (p. 50) for canceled tasks.

Type: Boolean

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException
The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

PipelineDeletedException
The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException
The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

TaskNotFoundException
The specified task was not found.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.ReportTaskProgress
Content-Length: 832
Host: datapipeline.us-east-1.amazonaws.com
```

```
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"taskId": "aaGgHT4LuH0T0Y0oLrJRjas5qH0d8cDPADxqq3tn+zCWGELk
CdV2JprLreXm1oxeP5EFZHFLJ69kjSsLYE0iYHYBYVGBrB+E/pYq7ANEEeGJFnSBMRiXZVA+8UJ3Oz
cInvXeinqBmBaKwii7hnnKb/AXjXiNTXyxgydX1KAyg1AxkwBYG4cfPYMZbuE
bQJFJvv5C/2+GVXz1w94nKYTeUeepwUOFOuRLS6JVtZoYwpF56E+Yfk1IcGpFO
vCZ01B4Bkuu7x3J+MD/j6kJgZLAgbCJQtI3eiW3kdGmX0p0I2BdY1ZsX6b4UiSvM3OMj6NE
HJCJL4E0ZfitnhCoe24Kvjo6C2hFbZq+ei/HPgSXBQM
Sagkr4vS9c0ChzxH2+LNYvec6bY4kymkaZI1dvOzmpa0FcnGf5AjSK4GpsV
iZ/ujz6zxFv81qBXzjF0/4M1775rjV1VUdyKaixiA/sJiACNezqZqETidp8d24BDPRhGsj6pBCrnelqG
Frk/gXEXUsJ+xwMifRC8UVwiKekpAvHUywVk7Ku4jH/n3i2VoLRP6FXwpUbelu34iiZ9czpXyLtyP
Kwxa87dlrnRVURwkcVjOt2Mcrcaqe+cbWHvNRhyrPkkdf
SF3ac8/wfgVbXvLEB2k9mKc67aD9rvdc1PKX09Tk8BKklsMT
pZ3TRCd4NzQlJKigMe8Jat9+1tKj4Ole5ZzW6uyTu2s2iFjEV8KXu4MaiRJyNKCdKeGh
hZWY37Qk4NBK4Ppgu+C6Y41dpfOh288SLDEVx0/UySlqOEd
hba7c6BiPp5r3hKj3mk9lFy5OYp1aoGLeeFmjXveTnPdf2gkWqXXg7AUbJ7jEs1F0lKZQg4szep2gcK
yAJXgvXLfJJHcha8Lfb/Ee7wYmyOcAaRpDBoFNSbtoVX
ar46teIrpho+ZDvynUXvU0grHWGOk=:wn3SgymHZM99bEXAMPLE",
 "fields":
  [
    {"key": "percentComplete",
     "stringValue": "50"}
  ]
}
```

Sample Response

```
x-amzn-RequestId: 640bd023-0775-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 18
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"canceled": false}
```

ReportTaskRunnerHeartbeat

Task runners call ReportTaskRunnerHeartbeat (p. 44) every 15 minutes to indicate that they are operational. In the case of AWS Data Pipeline Task Runner launched on a resource managed by AWS Data Pipeline, the web service can use this call to detect when the task runner application has failed and restart a new instance.

Request Syntax

```
{
    "Hostname": "string",
    "TaskrunnerId": "string",
    "WorkerGroup": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

Hostname
   The public DNS name of the calling task runner.

   Type: String

   Length constraints: Minimum length of 1. Maximum length of 1024.

   Required: No

TaskrunnerId
   The identifier of the task runner. This value should be unique across your AWS account. In the case of AWS Data Pipeline Task Runner launched on a resource managed by AWS Data Pipeline, the web service provides a unique identifier when it launches the application. If you have written a custom task runner, you should assign a unique identifier for the task runner.

   Type: String

   Length constraints: Minimum length of 1. Maximum length of 1024.

   Required: Yes

WorkerGroup
   Indicates the type of task the task runner is configured to accept and process. The worker group is set as a field on objects in the pipeline when they are created. You can only specify a single value for workerGroup in the call to ReportTaskRunnerHeartbeat (p. 44). There are no wildcard values permitted in workerGroup, the string must be an exact, case-sensitive, match.

   Type: String

   Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

Response Syntax

```
{
    "Terminate": "boolean"
}
```

Response Elements

The following data is returned in JSON format by the service.

Terminate

Indicates whether the calling task runner should terminate. If `True`, the task runner that called ReportTaskRunnerHeartbeat (p. 44) should terminate.

Type: Boolean

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError

An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException

The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.ReportTaskRunnerHeartbeat
Content-Length: 84
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"taskrunnerId": "1234567890",
 "workerGroup": "wg-12345",
```

```
"hostname": "example.com"}
```

Sample Response

```
Status:
x-amzn-RequestId: b3104dc5-0734-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 20
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"terminate": false}
```

SetStatus

Requests that the status of an array of physical or logical pipeline objects be updated in the pipeline. This update may not occur immediately, but is eventually consistent. The status that can be set depends on the type of object, e.g. DataNode or Activity. You cannot perform this operation on FINISHED pipelines and attempting to do so will return an InvalidRequestException.

Request Syntax

```
{
    "ObjectIds": [
        "string"
    ],
    "PipelineId": "string",
    "Status": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

ObjectIds
Identifies an array of objects. The corresponding objects can be either physical or components, but not a mix of both types.

Type: array of Strings

Required: Yes

PipelineId
Identifies the pipeline that contains the objects.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

Status
Specifies the status to be set on all the objects in `objectIds`. For components, this can be either `PAUSE` or `RESUME`. For instances, this can be either `CANCEL`, `RERUN`, or `MARK_FINISHED`.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError
> An internal service error occurred.
>
> HTTP Status Code: 500

InvalidRequestException
> The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.
>
> HTTP Status Code: 400

PipelineDeletedException
> The specified pipeline has been deleted.
>
> HTTP Status Code: 400

PipelineNotFoundException
> The specified pipeline was not found. Verify that you used the correct user and account identifiers.
>
> HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.SetStatus
Content-Length: 100
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-0634701J7KEXAMPLE",
 "objectIds":
  ["o-08600941GHJWMBR9E2"],
 "status": "pause"}
```

Sample Response

```
x-amzn-RequestId: e83b8ab7-076a-11e2-af6f-6bc7a6be60d9
Content-Type: application/x-amz-json-1.1
Content-Length: 0
Date: Mon, 12 Nov 2012 17:50:53 GMT
```

```
Unexpected response: 200, OK, undefined
```

SetTaskStatus

Notifies AWS Data Pipeline that a task is completed and provides information about the final status. The task runner calls this action regardless of whether the task was sucessful. The task runner does not need to call SetTaskStatus (p. 50) for tasks that are canceled by the web service during a call to ReportTaskProgress (p. 41).

Request Syntax

```
{
    "ErrorId": "string",
    "ErrorMessage": "string",
    "ErrorStackTrace": "string",
    "TaskId": "string",
    "TaskStatus": "string"
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

ErrorId

If an error occurred during the task, this value specifies an id value that represents the error. This value is set on the physical attempt object. It is used to display error information to the user. It should not start with string "Service_" which is reserved by the system.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

ErrorMessage

If an error occurred during the task, this value specifies a text description of the error. This value is set on the physical attempt object. It is used to display error information to the user. The web service does not parse this value.

Type: String

Required: No

ErrorStackTrace

If an error occurred during the task, this value specifies the stack trace associated with the error. This value is set on the physical attempt object. It is used to display error information to the user. The web service does not parse this value.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

TaskId

Identifies the task assigned to the task runner. This value is set in the TaskObject (p. 72) that is returned by the PollForTask (p. 27) action.

Type: String

Length constraints: Minimum length of 1. Maximum length of 2048.

Required: Yes

TaskStatus

If `FINISHED`, the task successfully completed. If `FAILED` the task ended unsuccessfully. The `FALSE` value is used by preconditions.

Type: String

Valid Values: `FINISHED` | `FAILED` | `FALSE`

Required: Yes

Response Elements

If the action is successful, the service sends back an HTTP 200 response with an empty HTTP body.

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError

An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException

The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

PipelineDeletedException

The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException

The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

TaskNotFoundException

The specified task was not found.

HTTP Status Code: 400

Examples

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.SetTaskStatus
Content-Length: 847
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"taskId": "aaGgHT4LuH0T0Y0oLrJRjas5qH0d8cDPADxqq3tn+zCWGELk
CdV2JprLreXm1oxeP5EFZHFLJ69kjSsLYE0iYHYBYVGBrB+E/pYq7ANEEeGJFnSBMRiXZVA+8UJ3Oz
cInvXeinqBmBaKwii7hnnKb/AXjXiNTXyxgydX1KAyg1AxkwBYG4cfPYMZbuE
bQJFJvv5C/2+GVXz1w94nKYTeUeepwUOFOuRLS6JVtZoYwpF56E+Yfk1IcGpFO
vCZ01B4Bkuu7x3J+MD/j6kJgZLAgbCJQtI3eiW3kdGmX0p0I2BdY1ZsX6b4UiSvM3OMj6NE
HJCJL4E0ZfitnhCoe24Kvjo6C2hFbZq+ei/HPgSXBQM
Sagkr4vS9c0ChzxH2+LNYvec6bY4kymkaZI1dvOzmpa0FcnGf5AjSK4GpsV
iZ/ujz6zxFv81qBXzjF0/4M1775rjV1VUdyKaixiA/sJiACNezqZqETidp8d24BDPRhGsj6pBCrnelqG
Frk/gXEXUsJ+xwMifRC8UVwiKekpAvHUywVk7Ku4jH/n3i2VoLRP6FXwpUbelu34iiZ9czpXyLtyP
Kwxa87dlrnRVURwkcVjOt2Mcrcaqe+cbWHvNRhyrPkkdf
SF3ac8/wfgVbXvLEB2k9mKc67aD9rvdc1PKX09Tk8BKklsMT
pZ3TRCd4NzQlJKigMe8Jat9+1tKj4Ole5ZzW6uyTu2s2iFjEV8KXu4MaiRJyNKCdKeGh
hZWY37Qk4NBK4Ppgu+C6Y41dpfOh288SLDEVx0/UySlqOEd
hba7c6BiPp5r3hKj3mk9lFy5OYp1aoGLeeFmjXveTnPdf2gkWqXXg7AUbJ7jEs1F0lKZQg4szep2gcK
yAJXgvXLfJJHcha8Lfb/Ee7wYmyOcAaRpDBoFNSbtoVX
ar46teIrpho+ZDvynUXvU0grHWGOk=:wn3SgymHZM99bEXAMPLE",
 "taskStatus": "FINISHED"}
```

Sample Response

```
x-amzn-RequestId: 8c8deb53-0788-11e2-af9c-6bc7a6be6qr8
Content-Type: application/x-amz-json-1.1
Content-Length: 0
Date: Mon, 12 Nov 2012 17:50:53 GMT

{}
```

ValidatePipelineDefinition

Tests the pipeline definition with a set of validation checks to ensure that it is well formed and can run without error.

Request Syntax

```
{
    "ParameterObjects": [
        {
            "Attributes": [
                {
                    "Key": "string",
                    "StringValue": "string"
                }
            ],
            "Id": "string"
        }
    ],
    "ParameterValues": [
        {
            "Id": "string",
            "StringValue": "string"
        }
    ],
    "PipelineId": "string",
    "PipelineObjects": [
        {
            "Fields": [
                {
                    "Key": "string",
                    "RefValue": "string",
                    "StringValue": "string"
                }
            ],
            "Id": "string",
            "Name": "string"
        }
    ]
}
```

Request Parameters

For information about the common parameters that all actions use, see Common Parameters (p. 75).

The request accepts the following data in JSON format.

ParameterObjects
A list of parameter objects used with the pipeline.

Type: array of ParameterObject (p. 65) objects

Required: No

ParameterValues

A list of parameter values used with the pipeline.

Type: array of ParameterValue (p. 66) objects

Required: No

PipelineId

Identifies the pipeline whose definition is to be validated.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

PipelineObjects

A list of objects that define the pipeline changes to validate against the pipeline.

Type: array of PipelineObject (p. 68) objects

Required: Yes

Response Syntax

```
{
    "Errored": "boolean",
    "ValidationErrors": [
        {
            "Errors": [
                "string"
            ],
            "Id": "string"
        }
    ],
    "ValidationWarnings": [
        {
            "Id": "string",
            "Warnings": [
                "string"
            ]
        }
    ]
}
```

Response Elements

The following data is returned in JSON format by the service.

Errored

If `True`, there were validation errors.

Type: Boolean

ValidationErrors

Lists the validation errors that were found by ValidatePipelineDefinition (p. 53).

Type: array of ValidationError (p. 73) objects

ValidationWarnings

Lists the validation warnings that were found by ValidatePipelineDefinition (p. 53).

Type: array of ValidationWarning (p. 74) objects

Errors

For information about the errors that are common to all actions, see Common Errors (p. 77).

InternalServiceError

An internal service error occurred.

HTTP Status Code: 500

InvalidRequestException

The request was not valid. Verify that your request was properly formatted, that the signature was generated with the correct credentials, and that you haven't exceeded any of the service limits for your account.

HTTP Status Code: 400

PipelineDeletedException

The specified pipeline has been deleted.

HTTP Status Code: 400

PipelineNotFoundException

The specified pipeline was not found. Verify that you used the correct user and account identifiers.

HTTP Status Code: 400

Examples

Example 1

This example sets an valid pipeline configuration and returns success.

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.ValidatePipelineDefinition
Content-Length: 936
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-06372391ZG65EXAMPLE",
 "pipelineObjects":
  [
    {"id": "Default",
     "name": "Default",
     "fields":
      [
```

```
                {"key": "workerGroup",
                 "stringValue": "MyworkerGroup"}
             ]
        },
        {"id": "Schedule",
         "name": "Schedule",
         "fields":
          [
            {"key": "startDateTime",
             "stringValue": "2012-09-25T17:00:00"},
            {"key": "type",
             "stringValue": "Schedule"},
            {"key": "period",
             "stringValue": "1 hour"},
            {"key": "endDateTime",
             "stringValue": "2012-09-25T18:00:00"}
          ]
        },
        {"id": "SayHello",
         "name": "SayHello",
         "fields":
          [
            {"key": "type",
             "stringValue": "ShellCommandActivity"},
            {"key": "command",
             "stringValue": "echo hello"},
            {"key": "parent",
             "refValue": "Default"},
            {"key": "schedule",
             "refValue": "Schedule"}

          ]
        }
     ]
}
```

Sample Response

```
x-amzn-RequestId: 92c9f347-0776-11e2-8a14-21bb8a1f50ef
Content-Type: application/x-amz-json-1.1
Content-Length: 18
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"errored": false}
```

Example 2

This example sets an invalid pipeline configuration and returns the associated set of validation errors.

Sample Request

```
POST / HTTP/1.1
Content-Type: application/x-amz-json-1.1
X-Amz-Target: DataPipeline.ValidatePipelineDefinition
Content-Length: 903
Host: datapipeline.us-east-1.amazonaws.com
X-Amz-Date: Mon, 12 Nov 2012 17:49:52 GMT
Authorization: AuthParams

{"pipelineId": "df-06372391ZG65EXAMPLE",
 "pipelineObjects":
  [
    {"id": "Default",
     "name": "Default",
     "fields":
      [
        {"key": "workerGroup",
         "stringValue": "MyworkerGroup"}
      ]
    },
    {"id": "Schedule",
     "name": "Schedule",
     "fields":
      [
        {"key": "startDateTime",
         "stringValue": "bad-time"},
        {"key": "type",
         "stringValue": "Schedule"},
        {"key": "period",
         "stringValue": "1 hour"},
        {"key": "endDateTime",
         "stringValue": "2012-09-25T18:00:00"}
      ]
    },
    {"id": "SayHello",
     "name": "SayHello",
     "fields":
      [
        {"key": "type",
         "stringValue": "ShellCommandActivity"},
        {"key": "command",
         "stringValue": "echo hello"},
        {"key": "parent",
         "refValue": "Default"},
        {"key": "schedule",
         "refValue": "Schedule"}

      ]
    }
  ]
}
```

Sample Response

```
x-amzn-RequestId: 496a1f5a-0e6a-11e2-a61c-bd6312c92ddd
Content-Type: application/x-amz-json-1.1
Content-Length: 278
Date: Mon, 12 Nov 2012 17:50:53 GMT

{"errored": true,
 "validationErrors":
  [
    {"errors":
      ["INVALID_FIELD_VALUE:  'startDateTime' value must be a literal datetime
 value."],
     "id": "Schedule"}
  ]
}
```

Data Types

The AWS Data Pipeline API contains several data types that various actions use. This section describes each data type in detail.

> **Note**
> The order of each element in the response is not guaranteed. Applications should not assume a particular order.

The following data types are supported:

ActivatePipelineResult

Description

Contains the output from the ActivatePipeline (p. 3) action.

Contents

CreatePipelineResult

Description

Contains the output from the CreatePipeline (p. 5) action.

Contents

PipelineId
> The ID that AWS Data Pipeline assigns the newly created pipeline. The ID is a string of the form: df-06372391ZG65EXAMPLE.

> Type: String

> Length constraints: Minimum length of 1. Maximum length of 1024.

> Required: Yes

DescribeObjectsResult

Description

If `True`, there are more results that can be returned in another call to DescribeObjects (p. 10).

Contents

HasMoreResults
> If `True`, there are more pages of results to return.

> Type: Boolean

> Required: No

Marker
> The starting point for the next page of results. To view the next page of results, call DescribeObjects (p. 10) again with this marker value.
>
> Type: String
>
> Length constraints: Minimum length of 0. Maximum length of 1024.
>
> Required: No

PipelineObjects
> An array of object definitions that are returned by the call to DescribeObjects (p. 10).
>
> Type: array of PipelineObject (p. 68) objects
>
> Required: Yes

DescribePipelinesResult

Description

Contains the output from the DescribePipelines (p. 14) action.

Contents

PipelineDescriptionList
> An array of descriptions returned for the specified pipelines.
>
> Type: array of PipelineDescription (p. 66) objects
>
> Required: Yes

EvaluateExpressionResult

Description

Contains the output from the EvaluateExpression (p. 17) action.

Contents

EvaluatedExpression
> The evaluated expression.
>
> Type: String
>
> Length constraints: Minimum length of 0. Maximum length of 20971520.
>
> Required: Yes

Field

Description

A key-value pair that describes a property of a pipeline object. The value is specified as either a string value (`StringValue`) or a reference to another object (`RefValue`) but not as both.

Contents

Key
> The field identifier.
>
> Type: String
>
> Length constraints: Minimum length of 1. Maximum length of 256.
>
> Required: Yes

RefValue
> The field value, expressed as the identifier of another object.
>
> Type: String
>
> Length constraints: Minimum length of 1. Maximum length of 256.
>
> Required: No

StringValue
> The field value, expressed as a String.
>
> Type: String
>
> Length constraints: Minimum length of 0. Maximum length of 10240.
>
> Required: No

GetPipelineDefinitionResult

Description

Contains the output from the GetPipelineDefinition (p. 20) action.

Contents

ParameterObjects
> Returns a list of parameter objects used in the pipeline definition.
>
> Type: array of ParameterObject (p. 65) objects
>
> Required: No

ParameterValues
> Returns a list of parameter values used in the pipeline definition.
>
> Type: array of ParameterValue (p. 66) objects

Required: No

PipelineObjects
An array of objects defined in the pipeline.

Type: array of PipelineObject (p. 68) objects

Required: No

InstanceIdentity

Description

Contents

Document
A description of an Amazon EC2 instance that is generated when the instance is launched and exposed to the instance via the instance metadata service in the form of a JSON representation of an object.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

Signature
A signature which can be used to verify the accuracy and authenticity of the information provided in the instance identity document.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

ListPipelinesResult

Description

Contains the output from the ListPipelines (p. 24) action.

Contents

HasMoreResults
If `True`, there are more results that can be obtained by a subsequent call to ListPipelines (p. 24).

Type: Boolean

Required: No

Marker
If not null, indicates the starting point for the set of pipeline identifiers that the next call to List-Pipelines (p. 24) will retrieve. If null, there are no more pipeline identifiers.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

PipelineIdList

A list of all the pipeline identifiers that your account has permission to access. If you require additional information about the pipelines, you can use these identifiers to call DescribePipelines (p. 14) and GetPipelineDefinition (p. 20).

Type: array of PipelineIdName (p. 67) objects

Required: Yes

Operator

Description

Contains a logical operation for comparing the value of a field with a specified value.

Contents

Type

The logical operation to be performed: equal (`EQ`), equal reference (`REF_EQ`), less than or equal (`LE`), greater than or equal (`GE`), or between (`BETWEEN`). Equal reference (`REF_EQ`) can be used only with reference fields. The other comparison types can be used only with String fields. The comparison types you can use apply only to certain object fields, as detailed below.

The comparison operators EQ and REF_EQ act on the following fields:

- name
- @sphere
- parent
- @componentParent
- @instanceParent
- @status
- @scheduledStartTime
- @scheduledEndTime
- @actualStartTime
- @actualEndTime

The comparison operators `GE`, `LE`, and `BETWEEN` act on the following fields:

- @scheduledStartTime
- @scheduledEndTime
- @actualStartTime
- @actualEndTime

Note that fields beginning with the at sign (@) are read-only and set by the web service. When you name fields, you should choose names containing only alpha-numeric values, as symbols may be reserved by AWS Data Pipeline. User-defined fields that you add to a pipeline should prefix their name with the string "my".

Type: String

Valid Values: `EQ` | `REF_EQ` | `LE` | `GE` | `BETWEEN`

Required: No

Values
   The value that the actual field value will be compared with.

   Type: array of Strings

   Required: No

ParameterAttribute

Description

The attributes allowed or specified with a parameter object.

Contents

Key
   The field identifier.

   Type: String

   Length constraints: Minimum length of 1. Maximum length of 256.

   Required: Yes

StringValue
   The field value, expressed as a String.

   Type: String

   Length constraints: Minimum length of 0. Maximum length of 10240.

   Required: Yes

ParameterObject

Description

Contains information about a parameter object.

Contents

Attributes
   The attributes of the parameter object.

   Type: array of ParameterAttribute (p. 65) objects

   Required: Yes

Id
   Identifier of the parameter object.

   Type: String

Length constraints: Minimum length of 1. Maximum length of 256.

Required: Yes

ParameterValue

Description

A value or list of parameter values.

Contents

Id
Identifier of the parameter value.

Type: String

Length constraints: Minimum length of 1. Maximum length of 256.

Required: Yes

StringValue
The field value, expressed as a String.

Type: String

Length constraints: Minimum length of 0. Maximum length of 10240.

Required: Yes

PipelineDescription

Description

Contains pipeline metadata.

Contents

Description
Description of the pipeline.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

Fields
A list of read-only fields that contain metadata about the pipeline: @userId, @accountId, and @pipelineState.

Type: array of Field (p. 62) objects

Required: Yes

Name

Name of the pipeline.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

PipelineId

The pipeline identifier that was assigned by AWS Data Pipeline. This is a string of the form `df-297EG78HU43EEXAMPLE`.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: Yes

Tags

Type: array of Tag (p. 71) objects

Length constraints: Minimum of 0 item(s) in the list. Maximum of 10 item(s) in the list.

Required: No

PipelineIdName

Description

Contains the name and identifier of a pipeline.

Contents

Id

Identifier of the pipeline that was assigned by AWS Data Pipeline. This is a string of the form `df-297EG78HU43EEXAMPLE`.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: No

Name

Name of the pipeline.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: No

PipelineObject

Description

Contains information about a pipeline object. This can be a logical, physical, or physical attempt pipeline object. The complete set of components of a pipeline defines the pipeline.

Contents

Fields
    Key-value pairs that define the properties of the object.

    Type: array of Field (p. 62) objects

    Required: Yes

Id
    Identifier of the object.

    Type: String

    Length constraints: Minimum length of 1. Maximum length of 1024.

    Required: Yes

Name
    Name of the object.

    Type: String

    Length constraints: Minimum length of 1. Maximum length of 1024.

    Required: Yes

PollForTaskResult

Description

Contains the output from the PollForTask action.

Contents

TaskObject
    An instance of PollForTaskResult (p. 68), which contains an instance of TaskObject (p. 72). The re-
    turned object contains all the information needed to complete the task that is being assigned to the
    task runner. One of the fields returned in this object is taskId, which contains an identifier for the task
    being assigned. The calling task runner uses taskId in subsequent calls to ReportTaskProgress (p. 41)
    and SetTaskStatus (p. 50).

    Type: TaskObject (p. 72) object

    Required: No

PutPipelineDefinitionResult

Description

Contains the output of the PutPipelineDefinition (p. 31) action.

Contents

Errored

If `True`, there were validation errors. If errored is `True`, the pipeline definition is stored but cannot be activated until you correct the pipeline and call PutPipelineDefinition (p. 31) to commit the corrected pipeline.

Type: Boolean

Required: Yes

ValidationErrors

A list of the validation errors that are associated with the objects defined in `pipelineObjects`.

Type: array of ValidationError (p. 73) objects

Required: No

ValidationWarnings

A list of the validation warnings that are associated with the objects defined in `pipelineObjects`.

Type: array of ValidationWarning (p. 74) objects

Required: No

Query

Description

Defines the query to run against an object.

Contents

Selectors

List of selectors that define the query. An object must satisfy all of the selectors to match the query.

Type: array of Selector (p. 71) objects

Required: No

QueryObjectsResult

Description

Contains the output from the QueryObjects (p. 37) action.

Contents

HasMoreResults

If `True`, there are more results that can be obtained by a subsequent call to QueryObjects (p. 37).

Type: Boolean

Required: No

Ids

A list of identifiers that match the query selectors.

Type: array of Strings

Required: No

Marker

The starting point for the results to be returned. As long as the action returns `HasMoreResults` as `True`, you can call QueryObjects (p. 37) again and pass the marker value from the response to retrieve the next set of results.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

ReportTaskProgressResult

Description

Contains the output from the ReportTaskProgress action.

Contents

Canceled

If `True`, the calling task runner should cancel processing of the task. The task runner does not need to call SetTaskStatus (p. 50) for canceled tasks.

Type: Boolean

Required: Yes

ReportTaskRunnerHeartbeatResult

Description

Contains the output from the ReportTaskRunnerHeartbeat (p. 44) action.

Contents

Terminate

Indicates whether the calling task runner should terminate. If `True`, the task runner that called Re-portTaskRunnerHeartbeat (p. 44) should terminate.

Type: Boolean

Required: Yes

Selector

Description

A comparision that is used to determine whether a query should return this object.

Contents

FieldName

The name of the field that the operator will be applied to. The field name is the "key" portion of the field definition in the pipeline definition syntax that is used by the AWS Data Pipeline API. If the field is not set on the object, the condition fails.

Type: String

Length constraints: Minimum length of 0. Maximum length of 1024.

Required: No

Operator

Contains a logical operation for comparing the value of a field with a specified value.

Type: Operator (p. 64) object

Required: No

SetTaskStatusResult

Description

The output from the SetTaskStatus (p. 50) action.

Contents

Tag

Description

No action documentation available.

Contents

Key
Type: String

Length constraints: Minimum length of 1. Maximum length of 128.

Required: Yes

Value
Type: String

Length constraints: Minimum length of 0. Maximum length of 256.

Required: Yes

TaskObject

Description

Contains information about a pipeline task that is assigned to a task runner.

Contents

AttemptId
Identifier of the pipeline task attempt object. AWS Data Pipeline uses this value to track how many times a task is attempted.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: No

Objects
Connection information for the location where the task runner will publish the output of the task.

Type: String to PipelineObject (p. 68) object map

Required: No

PipelineId
Identifier of the pipeline that provided the task.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: No

TaskId
An internal identifier for the task. This ID is passed to the SetTaskStatus (p. 50) and ReportTaskProgress (p. 41) actions.

Type: String

Length constraints: Minimum length of 1. Maximum length of 2048.

Required: No

ValidatePipelineDefinitionResult

Description

Contains the output from the ValidatePipelineDefinition (p. 53) action.

Contents

Errored
> If `True`, there were validation errors.
>
> Type: Boolean
>
> Required: Yes

ValidationErrors
> Lists the validation errors that were found by ValidatePipelineDefinition (p. 53).
>
> Type: array of ValidationError (p. 73) objects
>
> Required: No

ValidationWarnings
> Lists the validation warnings that were found by ValidatePipelineDefinition (p. 53).
>
> Type: array of ValidationWarning (p. 74) objects
>
> Required: No

ValidationError

Description

Defines a validation error returned by PutPipelineDefinition (p. 31) or ValidatePipelineDefinition (p. 53). Validation errors prevent pipeline activation. The set of validation errors that can be returned are defined by AWS Data Pipeline.

Contents

Errors
> A description of the validation error.
>
> Type: array of Strings
>
> Required: No

Id
> The identifier of the object that contains the validation error.
>
> Type: String
>
> Length constraints: Minimum length of 1. Maximum length of 1024.
>
> Required: No

ValidationWarning

Description

Defines a validation warning returned by PutPipelineDefinition (p. 31) or ValidatePipelineDefinition (p. 53). Validation warnings do not prevent pipeline activation. The set of validation warnings that can be returned are defined by AWS Data Pipeline.

Contents

Id

The identifier of the object that contains the validation warning.

Type: String

Length constraints: Minimum length of 1. Maximum length of 1024.

Required: No

Warnings

A description of the validation warning.

Type: array of Strings

Required: No

Common Parameters

This section lists the request parameters that all actions use. Any action-specific parameters are listed in the topic for the action.

Action
The action to be performed.

Default: None

Type: string

Required: Yes

AuthParams
The parameters that are required to authenticate a Conditional request. Contains:

- AWSAccessKeyID
- SignatureVersion
- Timestamp
- Signature

Default: None

Required: Conditional

AWSAccessKeyId
The access key ID that corresponds to the secret access key that you used to sign the request.

Default: None

Type: string

Required: Yes

Expires
The date and time when the request signature expires, expressed in the format YYYY-MM-DDThh:mm:ssZ, as specified in the ISO 8601 standard.

Condition: Requests must include either *Timestamp* or *Expires*, but not both.

Default: None

Type: string

Required: Conditional

SecurityToken

The temporary security token that was obtained through a call to AWS Security Token Service. For a list of services that support AWS Security Token Service, go to Using Temporary Security Credentials to Access AWS in **Using Temporary Security Credentials**.

Default: None

Type: string

Required: No

Signature

The digital signature that you created for the request. For information about generating a signature, go to the service's developer documentation.

Default: None

Type: string

Required: Yes

SignatureMethod

The hash algorithm that you used to create the request signature.

Default: None

Type: string

Valid Values: `HmacSHA256 | HmacSHA1`

Required: Yes

SignatureVersion

The signature version you use to sign the request. Set this to the value that is recommended for your service.

Default: None

Type: string

Required: Yes

Timestamp

The date and time when the request was signed, expressed in the format YYYY-MM-DDThh:mm:ssZ, as specified in the ISO 8601 standard.

Condition: Requests must include either *Timestamp* or *Expires*, but not both.

Default: None

Type: string

Required: Conditional

Version

The API version that the request is written for, expressed in the format YYYY-MM-DD.

Default: None

Type: string

Required: Yes

Common Errors

This section lists the common errors that all actions return. Any action-specific errors are listed in the topic for the action.

IncompleteSignature
> The request signature does not conform to AWS standards.
>
> HTTP Status Code: 400

InternalFailure
> The request processing has failed because of an unknown error, exception or failure.
>
> HTTP Status Code: 500

InvalidAction
> The action or operation requested is invalid. Verify that the action is typed correctly.
>
> HTTP Status Code: 400

InvalidClientTokenId
> The X.509 certificate or AWS access key ID provided does not exist in our records.
>
> HTTP Status Code: 403

InvalidParameterCombination
> Parameters that must not be used together were used together.
>
> HTTP Status Code: 400

InvalidParameterValue
> An invalid or out-of-range value was supplied for the input parameter.
>
> HTTP Status Code: 400

InvalidQueryParameter
> The AWS query string is malformed or does not adhere to AWS standards.
>
> HTTP Status Code: 400

MalformedQueryString
> The query string contains a syntax error.
>
> HTTP Status Code: 404

MissingAction
> The request is missing an action or a required parameter.

HTTP Status Code: 400

MissingAuthenticationToken

The request must contain either a valid (registered) AWS access key ID or X.509 certificate.

HTTP Status Code: 403

MissingParameter

A required parameter for the specified action is not supplied.

HTTP Status Code: 400

OptInRequired

The AWS access key ID needs a subscription for the service.

HTTP Status Code: 403

RequestExpired

The request reached the service more than 15 minutes after the date stamp on the request or more than 15 minutes after the request expiration date (such as for pre-signed URLs), or the date stamp on the request is more than 15 minutes in the future.

HTTP Status Code: 400

ServiceUnavailable

The request has failed due to a temporary failure of the server.

HTTP Status Code: 503

Throttling

The request was denied due to request throttling.

HTTP Status Code: 400

ValidationError

The input fails to satisfy the constraints specified by an AWS service.

HTTP Status Code: 400